

WU, F., WANG, S., LIU, D., CAO, W., FERNANDEZ, C. and HUANG, Q. 2024. An improved convolutional neural network-bidirectional gated recurrent unit algorithm for robust state of charge and state of energy estimation of new energy vehicles of lithium-ion batteries. *Journal of energy storage* [online], 82, article 110574. Available from: <https://doi.org/10.1016/j.est.2024.110574>

An improved convolutional neural network-bidirectional gated recurrent unit algorithm for robust state of charge and state of energy estimation of new energy vehicles of lithium-ion batteries.

WU, F., WANG, S., LIU, D., CAO, W., FERNANDEZ, C. and HUANG, Q.

2024

An improved convolutional neural network-bidirectional gated recurrent unit algorithm for robust state of charge and state of energy estimation of new energy vehicles of lithium-ion batteries

Fan Wu^a, Shunli Wang^{a*,b}, Donglei Liu^a, Wen Cao^a, Carlos Fernandez^c, Qi Huang^a

^a*School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010,*

China; ^b*College of Electrical Engineering, Sichuan University, Chengdu 610065, China;* ^c*School of*

Pharmacy and Life Sciences, Robert Gordon University, Aberdeen AB10-7GJ, UK.

Abstract: State of charge (SOC) and state of energy (SOE) are the key factors that reflect the safe and range driving of new energy vehicles. This paper proposes an optimized convolutional neural network-bidirectional gate recurrent unit (CNN-BiGRU) and an improved Kalman bidirectional smoothing algorithm to predict SOC and SOE accurately. Firstly, the attention mechanism (AM) is introduced into the CNN-BiGRU to extract significant features better. A multi-task learning (MTL) mechanism is constructed to learn the correlation between tasks and output the results simultaneously. Then, square root and reverse smoothing methods are added to the classical extended Kalman filtering (SREKS) to filter and de-noise the neural network output. Finally, the test data of the hybrid pulse power characterization (HPPC) and Beijing bus dynamic stress test (BBDST) at 15□ and 35□ are used for experimental verification. Under the HPPC condition, the maximum error of SOC and SOE is less than 0.01432 and 0.01417, respectively. Under BBDST condition, the maximum error of SOC and SOE is less than 0.01827 and 0.01883, respectively. Experimental results show that this algorithm has high accuracy and robustness under different complex working conditions.

Keywords: convolutional neural network; bidirectional gate recurrent unit; attention mechanism; multi-task learning mechanism; square root extended Kalman filter; lithium-ion battery

Corresponding author: *Shunli Wang. Tel/fax: +86-15884655563. E-mail address: wangshunli@swust.edu.cn.*

1. Introduction

With the massive development and utilization of fossil fuels in our country, the problems of resource shortage and environmental pollution have become increasingly prominent. New energy vehicles (NEVs) driven by batteries are the direction of development in the automotive field. Lithium-ion batteries are widely used as power sources for NEVs due to their long cycle life, low self-discharge rate, and high energy density [1, 2]. However, lithium-ion battery explosion accidents have occurred frequently at home and abroad in recent years. Therefore, a good battery management system (BMS) is essential for lithium-ion batteries' efficiency and safety operation [3, 4]. The accurate estimation of SOC and SOE in BMS can further ensure the efficient operation of the battery system. SOC represents the battery's remaining capacity, and SOE represents the battery's remaining energy [5, 6]. At the same time, the unbiased estimation of SOC and SOE can further rationally allocate power and extend battery range.

SOC is expressed as a ratio of the remaining available capacity over the nominal capacity. SOE is expressed as a ratio of the battery residual energy under specific operating conditions. The common SOC and SOE estimation methods are based on the characteristic parameter (CP), ampere-hour integration (AHI), battery equivalent circuit model (ECM), and data-driven [7-10]. The open circuit voltage (OCV) method of CP adopts the nonlinear correspondence between voltage and SOE or SOC to achieve the estimation purpose. Such as, Wang et al. [11] proposed cloud-based data to construct OCV-SOC complete curve method at different temperatures. Although the OCV method can directly reflect SOC changes, the battery needs to undergo long-term quiescence processing, which makes the prediction time longer. The AHI method calculates SOC or SOE by integrating voltage and current over time. Liu et al. [12] proposed a new method for SOC fuzzy control strategy based on an extended Kalman filtering (KF) and ampere-hour. Although AHI can quickly obtain an estimate from a mathematical equation, the initial value directly affects the accuracy of estimates and the integration operation results in the

accumulation of errors. The method based on the ECM has mainly used the KF. Lai et al. [13] proposed a joint SOH and SOE estimation method for lithium-ion batteries by combining the forgetting factor recursive least squares (FFRLS) and unscented KF. However, the KF algorithm depends on the accurate construction of the ECM, which is bound to consume much time and be challenging.

The emergence of the data-driven method dramatically reduces the time required for prediction and the complexity caused by the accuracy of ECM building. Machine learning and deep learning are the two main data-driven approaches [14, 15]. Basic machine learning algorithms include support vector machine, linear regression, and k-nearest neighbors [16-18]. It is precise because of the simple structure that combining complex electrochemical, electrothermal, and degradation dynamic behaviors is challenging under extremely complex working conditions [19]. Deep learning is more suitable for estimating BMS-related performance indicators than machine learning and is widely used [20, 21]. Deep learning includes generative adversarial networks (GAN), convolutional neural networks (CNN), and recurrent neural networks (RNN) [22-25]. RNN is widely used in the prediction of time series data because of their memory properties. Zhang et al. [26] proposed an improved method of BiGRU networks based on Nesterov Accelerated Gradient (NAG) algorithm, and experimental results show that the method can improve the prediction accuracy at different ambient temperatures.

However, because of the problem of gradient disappearance and gradient explosion of RNN, and the problem of long-term dependence, the LSTM network is proposed [27, 28]. LSTM can use different battery data to train corresponding network parameter models, which have certain flexibility [29]. Ma et al. [30] investigated a data-driven approach based on an improved LSTM for simultaneous estimation of SOC and SOE. As an improved network of LSTM, the gated recursive unit (GRU) is optimized from three to two gated units. As a result, GRU have a much simpler structure, allowing for fewer model parameters and faster computation. In addition, literature [31] shows that the prediction speed of GRU is 22.79% faster than LSTM, and experimental results of

literature [32-35] all indicate that the prediction effect of GRU model is better than LSTM. Dong et al. [36] established BiGRU circuit module, attention circuit module and KF module, formed the strategy of efficient information extraction and accurate filtering, and realized the high precision SOC estimation.

Compared with the single-state estimate, the joint estimation of SOC and SOE can better reflect the battery power condition under the condition of sudden voltage change. It provides a firmer basis for the functions of energy distribution, balance control, and charge and discharge control in the BMS so as to effectively extend the battery mileage and service life. In addition, a multi-task simultaneous estimation can effectively avoid the overfitting problem of a single-task estimation model. At present, most of the research on SOC and SOE joint estimation is based on ECM, which obtains state value through an improved parameter identification method combined with the KF algorithm [37-39]. However, with the continuous aging of the battery, the ECM model has difficulty expressing the complex changes inside the battery, which makes the prediction accuracy low[40]. In the estimation of SOC and SOE based on neural networks, as in the literature [41, 42]. In the simultaneous estimation of SOC and SOE, they ignore the coupling relationship between SOC and SOE. Almost all of them use the LSTM network, ignoring the GRU, which is optimized on the basis of LSTM. GRU has fewer model parameters and less computational effort.

In addition, to obtain a more accurate estimate, a filter method is introduced after the operation of the neural network, and the data is smoothed and denoised [43]. For example, Pang et al. [44] combined a backpropagation (BP) neural network and an extended Kalman particle filter (EKPF) algorithm to comprehensively estimate the SOC, the experimental results prove that the combined method based on the neural network has higher accuracy and robustness than (particle filter) PF, EKF, and EKPF. The joint of the neural network and KF method does not need to establish the ECM of the lithium-ion battery but only needs a certain amount of original data to train the network model to obtain accurate estimation results.

Based on the above problems, to obtain high precision estimates of SOC and SOE, an improved convolutional bidirectional gate recurrent unit-square root extended Kalman smoothing algorithm (AMC-BiGRU-SREKS) is proposed for simultaneous estimation of SOC and SOE. The main contributions of this paper are as follows.

- (1) Based on the solid nonlinearity and sequential nature of battery data, CNN-BiGRU model based on multi-sliding window data preprocessing is constructed. It can learn nonlinear relationships quickly and capture feature information bidirectionally, which effectively improves the estimation accuracy and speed.
- (2) A Luong attention model based on the global attention mechanism is constructed to assign different weights to the hidden state of CNN-BiGRU to increase the influence of important information on SOC and SOE.
- (3) According to the coupling relationship between SOC and SOE, a multi-task learning framework based on a hard sharing mechanism is constructed to form a new method of joint estimation for information sharing and complementarity between multiple tasks.
- (4) A high-precision bidirectional filtering strategy is constructed to address the problem of model accuracy caused by noise interference. The forward is to introduce a square root algorithm on the basis of EKF to ensure filtering performance. The fixed-interval smoothing technique is introduced backward to achieve high-precision joint prediction of SOC and SOE.
- (5) The accuracy of the novel AMC-BiGRU-SREKS method is verified under HPPC and BBDST working conditions at different temperatures. In addition, the mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean square error (RMSE) values of this algorithm are significantly better than those of other algorithms.

2. Principle analysis and model construction

2.1. Improved convolutional neural network-bidirectional gated recurrent unit

2.1.1. Convolutional neural network

Sample data is the key to constructing a neural network prediction model. In view of the strong nonlinearity and timing properties of the battery sample data, this paper adopts multi-window sliding window technology to process the sample data so as to characterize the data properties effectively. The schematic diagram of multi-sliding window construction sample data is shown in Fig. 1. Where the total period is T , the sliding window $w=3$, and new sample data is obtained by sliding one time step along the time axis each time, and a total of $T-w+1$ sample data can be constructed.

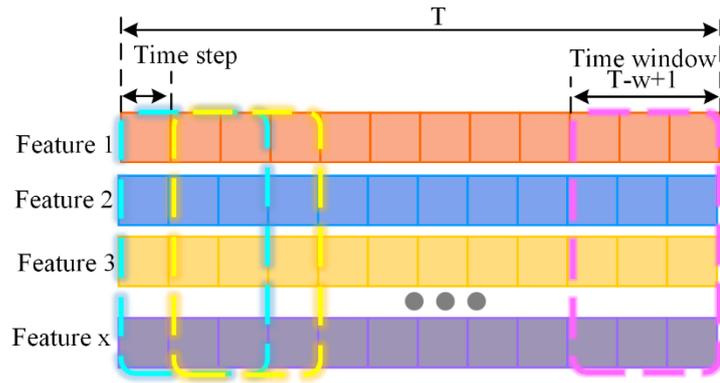


Fig. 1. Multi-sliding window data processing

As a feedforward neural network, CNN is often used in feature extraction. The most basic CNN is composed of the input, convolutional, pooling, full connection, and output layers. The network distribution of CNN is shown in Fig. 2. In the convolutional layer, convolution operation is used to extract the features of input data. Considering the strong nonlinearity of battery parameters and the linearity of a single convolutional process, RELU or other activation functions are introduced after convolutional operation to improve the nonlinear characterization ability of the network. The pooling layer integrates features from the convolutional transport layer and uses the maximum pooling layer to select the maximum eigenvalues and reduce the dimensions. The

fully connected layer integrates the data features from the pooled layer and uses appropriate activation functions for nonlinear transformation. It allows the network to gradually understand the complex features of the input data to make more accurate predictions.

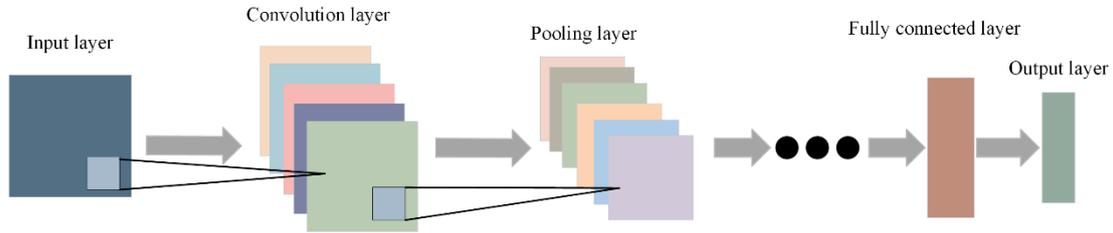


Fig. 2. Structure of CNN

2.1.2. Bidirectional gated recurrent unit

In contrast to the three-door structure of LSTM, the GRU combines the forgot door and the input door into a single update door. Thus, the GRU contains only an updated door and a reset door. The update gate defines how many steps the previous memory is saved to the current time. The reset gate determines how new input information is combined with the previous memory. The proposed GRU structure not only speeds up the convergence rate but also avoids the gradient problems in RNN. The internal structure of the GRU and LSTM are shown in Fig. 3.

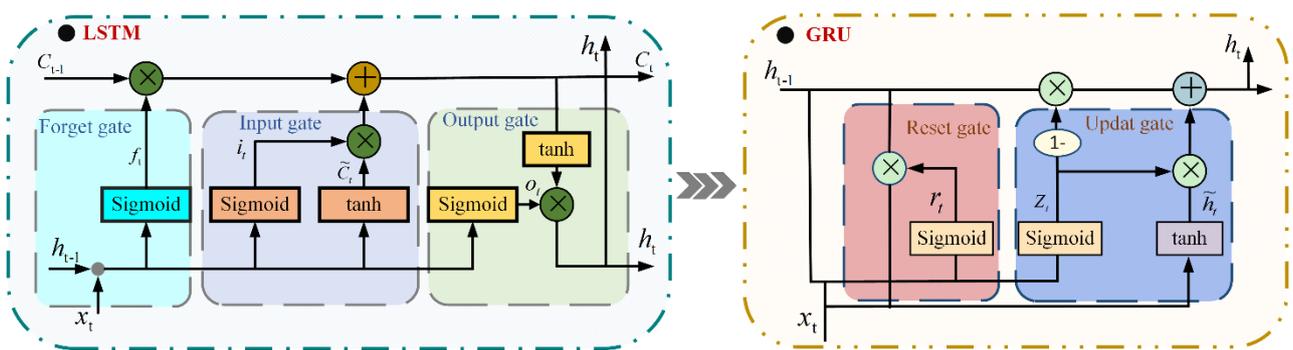


Fig. 3. Internal structure diagram of the GRU and LSTM

In Fig. 2, It is obvious that the GRU has a simpler internal structure than the LSTM. Where, h_{t-1} is the state transmitted from the previous one, x_t is the input of the current node, h_t is the output of the hidden layer, r_t is the reset gate, z_t is the update gate, \tilde{h}_t is the candidate's hidden state, and 1- represents the data transmitted

forward by the link as $1-z_t$. The sigmoid function can transform the data into a value in the range of 0 to 1, thus acting as a gated signal. When r_t approaches zero, the model will keep only the current input information and delete previously hidden information. When r_t approaches 1, the model will retain information from the past. z_t ranges from 0 to 1. The closer the value of the gated signal is to 1, the more information is saved. The units based on GRU can be calculated by Equation (1).

$$\begin{cases} z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \\ r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \\ \tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \\ h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \end{cases} \quad (1)$$

Where, W_z , W_r and W are trainable parameter matrices. The classical GRU structure propagates unidirectionally along the direction of sequence transition, with time t only related to past time. However, In the battery power from high to the end of a cycle, the SOC and SOE values will show a decreasing trend with the battery use time. When using the time series properties of SOC and SOE for estimation, the influence of past and future data on the current value can be considered to achieve accurate estimates. Therefore, establishing a forward and reverse BiGRU model is conducive to learning the relationship between past and future factors influencing battery and current state parameters. The structure of BiGRU is shown in Fig. 4.

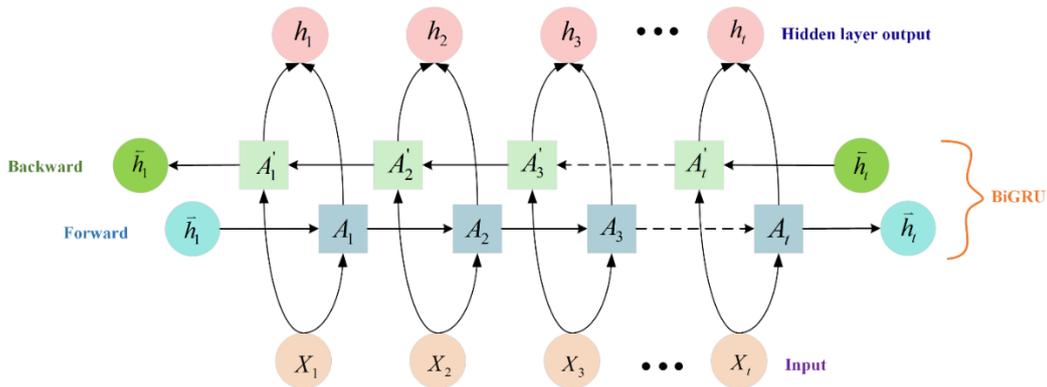


Fig. 4. Structure diagram of BiGRU

2.1.3. Attention mechanism

In neural network learning, increasing the complexity of the model is usually used to enhance the

expressiveness of the model. However, with the increase in the complexity of the model, the problem of information overload will often appear. Therefore, this paper introduces the Luong attention model based on global attention, assigns corresponding weights to all hidden states, strengthens the influence of important information on output, and ignores irrelevant information. Compared with the Bahdanau attention, which only considers partially hidden states, the Luong attention calculation logic is more apparent, which can effectively solve the problem of information overload and improve the efficiency and accuracy of the overall model. The structure of the improved AM is shown in Fig. 5.

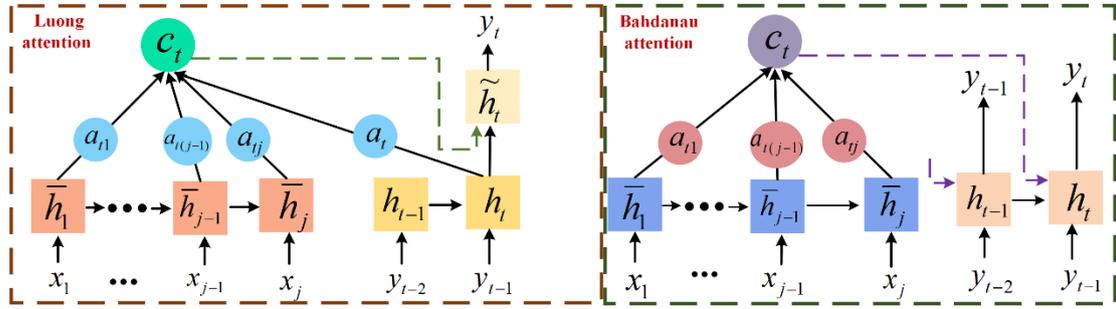


Fig. 5. The structure of the improved AM

In Fig. 5, \bar{h}_j is the hidden state at the input end, h_t is the hidden state at the output end, a_t is the attention weight, C_t is the context vector, \tilde{h}_t is the state of the current hidden layer with the attention mechanism added, x_j is the input of the CNN-BiGRU hidden layer at time j , and y_t is the output of AM at time t . It can be seen from the figure that Luong attention includes all the current hidden states in the calculation of attention, while Bahdanau attention does not consider the current hidden layer state. In contrast, Luong attention contains more feature information and has more rigorous logic. The calculation process of Luong attention is shown in Equation 2.

$$\left\{ \begin{array}{l} e_{tj} = \text{utanh}(w_a[h_t; \bar{h}_j]) \\ a_{tj} = \text{align}(h_t, \bar{h}_j) = \frac{\exp(e_{tj})}{\sum_{j=1}^T \exp(e_{tj})} \\ c_t = \sum_{j=1}^T a_{tj} \bar{h}_j \\ h_t = f(h_{t-1}, y_{t-1}) \\ \tilde{h}_t = \text{tanh}(w_c[c_t; h_t]) \end{array} \right. \quad (2)$$

Where, e_{tj} is the alignment fraction of input sequence time j to output sequence time t , and f is CNN-BiGRU cell unit. u , w_a and w_c are weight coefficients. When $t=n$, the current output hidden state h_n and each input hidden state \bar{h}_n are calculated using the concat function to obtain the alignment score e_{nj} , and the attention weight a_{nj} is obtained using the align function. Finally, the current hidden state with attention is obtained according to c_n and h_n , and the output y_n is obtained. Through the continuous update of a_{tj} , the final accurate estimate is obtained.

2.1.4. Multi-task learning mechanism

In BMS, the reliability analysis of the battery's SOC, SOE, SOH, SOP (state of power), and SOS (state of safety) is a prerequisite for the management of battery charging and discharge efficiency, safety, and health management [45]. However, most kinds of literature predict a single state parameter of the battery, leading to an inaccurate judgment of the battery-related performance. In dealing with a problem, the sharing layer can be used to learn and obtain the auxiliary coupling information provided by other related subtasks. Compared with single-task learning (STL), the multi-task learning (MTL) has a better effect, more vital generalization ability, and broader application scope. MTL contains a set of task $y^t (t \in T)$ and data set $\{x_i, y_i^1, y_i^2, \dots, y_i^t, \dots, y_i^T\} (i = 1, 2, \dots, N)$, where T is the estimated number of tasks, N is the sample data, and y_i^t is the label of the t th task at the i th data point. The prediction function is defined as $f^t(x; \theta^{sh}, \theta^t): x \rightarrow y^t$, where θ^{sh} is the parameter shared by different tasks, θ^t is the parameter related to tasks. Then the overall optimization loss function of MTL can be defined as shown in Equation (3).

$$\min \sum_{t=1}^T \alpha^t \hat{L}^t(\theta^{sh}, \theta^t) \quad (3)$$

Where, α^t is the weight coefficient of the task, $\hat{L}^t(\theta^{sh}, \theta^t)$ is the loss function. The optimization goal of the model is to minimize the task loss. MSE is used as the loss function and Adam is used as the parameter optimizer.

Two commonly used methods in MTL based on deep NN are hard sharing and soft sharing of hidden layer

parameters. The hard-sharing of parameters uses the same feature-sharing layer for multiple subtasks. It can effectively reduce the risk of overfitting through average noise. The parameter soft-sharing mechanism uses different network and characteristic parameters for different tasks. Compared with the soft sharing mechanism, the hard sharing mechanism builds a simpler model, and the sharing mechanism is more suitable for this study. The complex shared MTL of hidden layer parameters estimates SOE and SOC values. Under the condition of sharing information such as voltage, current, and temperature, relevant information on the coupling relationship between tasks can be found, which not only enhances the estimation accuracy of the model but also speeds up the estimation efficiency. The structure of the hard-sharing mechanism for MTL is shown in Fig. 6.

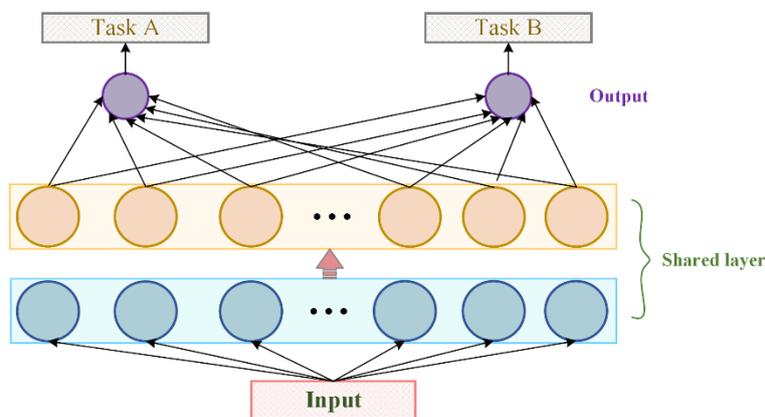


Fig. 6. The structure of the hard-sharing mechanism for multi-tasking learning

2.1.5. Overall optimization of the BiGRU algorithm

In the AMC-BiGRU model, constructing BiGRU can solve the timing problem of forecast data well and strengthen the time correlation between data. To improve the efficiency of data processing by BiGRU, the CNN-BiGRU model is constructed by introducing CNN to extract important feature information before BiGRU. Due to the large and complex input data, AM is introduced to give BiGRU different probability weights of hidden states and focus on more effective messages. A multi-task attention layer is constructed to make the model focus on the important and valid information of different tasks. The construction of the AMC-BiGRU model can better extract features, solve the long-term dependence problem between time series data, more accurately focus on the

key factors affecting the output, better learn the coupling relationship between the output. The structural model of AMC-BiGRU is shown in Fig. 7.

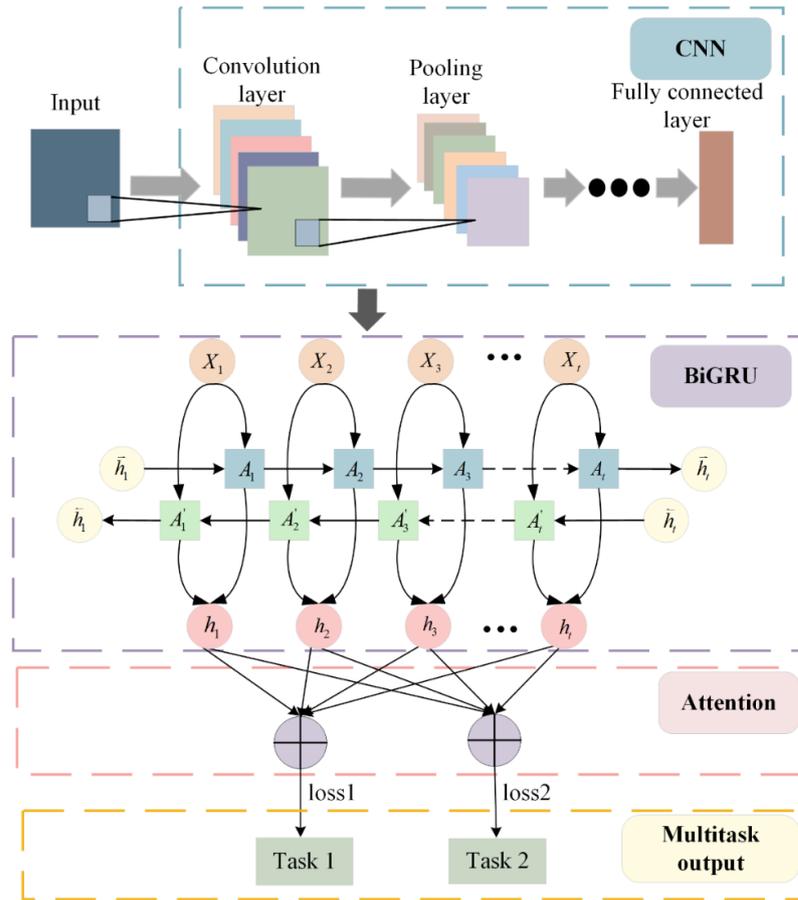


Fig. 7. Structure diagram of AMC-BiGRU model

2.2. Bidirectional smoothing filter

Although the AMC-BiGRU neural network algorithm can use the original battery data as input and output relatively accurate estimates, it will also bring some estimation errors due to the setting of hyperparameters and noise interference in practical applications. To increase the robustness of the model, a bidirectional denoising SREKS algorithm is constructed. It is divided into forward filtering and backward smoothing, and the SOC and SOE numerical curves from AMC-BiGRU are optimized to obtain more accurate estimates.

1) Forward filtering

SREKF has the same state-space equation as EKF, is shown in Equation (4).

$$\begin{cases} x_{k+1} = f(x_k, u_k) + w_k = A_k x_k + B_k u_k + w_k \\ Z_k = g(x_k, u_k) + v_k = C_k x_k + D_k u_k + v_k \end{cases} \quad (4)$$

Where, $f(x_k, u_k)$ and $g(x_k, u_k)$ are nonlinear functions in the system respectively, and state space equations of the system are formed by introducing state noise and observation noise. x_k is the n-dimensional state variable of the system, A_k is the state transition matrix of the system, B_k is the control input matrix of the system, u_k is the system input variable, w_k process measurement noise sequence, Z_k is the observation variable, C_k is the observation matrix of the system, D_k is the feedforward matrix of the system, v_k is the observe the noise sequence. The derivation formula of SREKF algorithm is shown as follows.

First, the error covariance matrix is decomposed into the lower triangular matrix S_k and the transpose S_k^T of the lower triangular matrix by Cholesky:

$$\begin{cases} P_k = S_k S_k^T \\ P'_{k+1} = A_k P_k A_k^T = A_k S_k S_k^T A_k^T = \tilde{S}'_{k+1} (\tilde{S}'_{k+1})^T \\ \tilde{S}'_{k+1} = A_k S_k v \\ P_{k+1} = \tilde{S}'_{k+1} (\tilde{S}'_{k+1})^T - \tilde{S}'_{k+1} (\tilde{S}'_{k+1})^T C_k^T C_k \tilde{S}'_{k+1} (\tilde{S}'_{k+1})^T + R_k C_k \tilde{S}'_{k+1} (\tilde{S}'_{k+1})^T \end{cases} \quad (5)$$

If $F_k = (\tilde{S}'_{k+1})^T C_k^T$ is satisfied, then it can be proven that :

$$\begin{cases} P_{k+1} = \tilde{S}'_{k+1} \{E - F_k [F_k^T F_k + R_k]^{-1} (\tilde{S}'_{k+1})^T\} = \tilde{S}'_{k+1} [E - \alpha_k F_k^T F_k (\tilde{S}'_{k+1})^T] (\tilde{S}'_{k+1})^T \\ \alpha_k = [F_k^T F_k + R_k]^{-1} \end{cases} \quad (6)$$

If $P_{k+1} = \tilde{S}'_{k+1} \tilde{S}'_{k+1}$ is satisfied, then it can be shown that:

$$\begin{cases} E - \alpha_k F_k^T F_k = E - \alpha_k F_k [2\gamma_k F_k^T - \alpha_k \gamma_k^2 F_k^T F_k] F_k^T \\ \tilde{S}'_{k+1} = \tilde{S}'_{k+1} [E - \alpha_k \gamma_k F_k F_k^T] \end{cases} \quad (7)$$

According to Equation (5)~(7), the SREKF algorithm can be obtained as follow:

$$\begin{cases} x'_k = f(x_{k-1}, u_{k-1}) \\ \tilde{S}'_k = A_{k-1} S_{k-1} \\ K_{k-1} = \alpha_{k-1} \tilde{S}'_{k-1} F_{k-1} \\ \alpha_{k-1} = [F_{k-1}^T F_{k-1} + R_{k-1}]^{-1} \\ F_{k-1} = (\tilde{S}'_{k-1})^T C_{k-1}^T \\ x_k = x'_k + K_{k-1} [Z_{k-1} - A_{k-1} C_{k-1} x'_k] \\ \tilde{S}'_k = \tilde{S}'_{k-1} [E - \alpha_{k-1} \gamma_{k-1} F_{k-1} F_{k-1}^T] \\ \gamma_{k-1} = \frac{1 \pm \sqrt{\alpha_{k-1} R_{k-1}}}{1 - \alpha_{k-1} R_{k-1}} \end{cases} \quad (8)$$

2) Reverse smoothing

The reverse smoothing algorithm smoothes the result based on forward filtering, reducing fluctuation amplitude

to obtain stable output results and lower error. The reverse smoothing is shown in Equation (9).

$$\begin{cases} x'_{k+1} = A_k x_k \\ \tilde{S}'_{k+1} = A_k \tilde{S}_k A_k^T + Q_k \\ G_k = \tilde{S}_k A_k^T (S'_{k+1})^{-1} \\ x_k^S = x_k + G_k (x'_{k+1} - x'_{k+1}) \\ \tilde{S}_k^S = S_k + G_k (\tilde{S}'_{k+1} - \tilde{S}'_{k+1}) G_k^T \end{cases} \quad (9)$$

Based on the forward calculation of the N-dimensional state variable x_k and covariance matrix S_k . A new Kalman gain G_k , state variable x_k^S and covariance matrix \tilde{S}_k^S are obtained. Finally, the results of the AMC-BiGRU-SREKF algorithm are reverse-smoothed, which improves the robustness of the system and the accuracy of the model.

2.3. The overall framework of the AMC-BiGRU-SREKS model

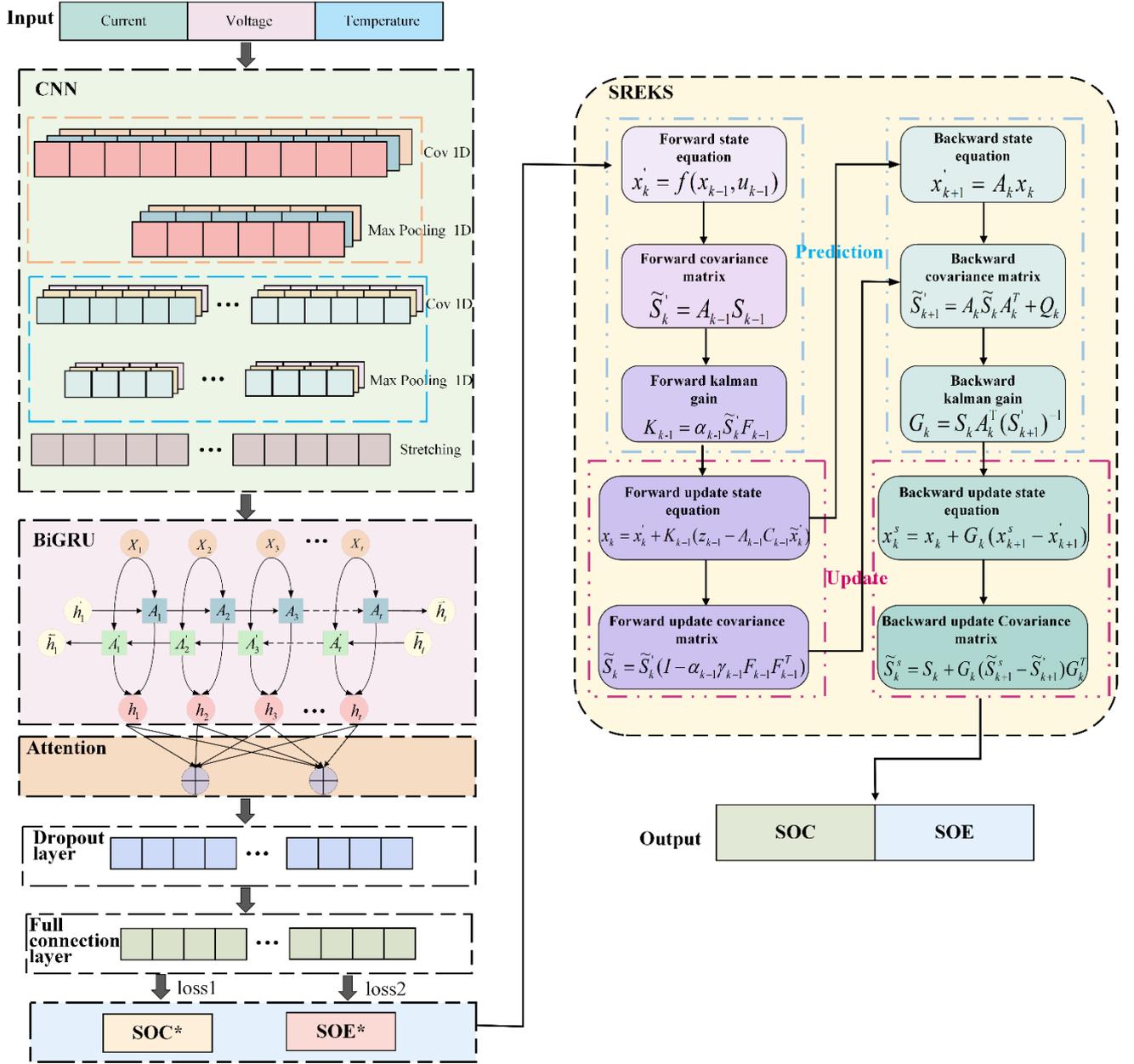


Fig. 8. Structure of the AMC-BiGRU-SREKS model

Firstly, the combined voltage, current, and temperature data after sliding window preprocessing are used as input for one-dimensional CNN. CNN also carries out the convolution operation of multiple convolution kernels for the whole input sequence data in the form of a sliding window. The default step size of the sliding window is 1, and the window is the convolution kernel size. The same time step contains voltage, current and temperature data. After the convolution operation, the nonlinear mapping of the activation function and dimensionality

reduction of the pooling layer is carried out successively on the feature data. In order to extract high-level feature data, two convolution layers and pooling layers are used in this paper. The final step in CNN is to expand the data from the second pooled layer into one-dimensional data and make it as input to BiGRU. BiGRU uses two directional GRU to learn the feature information extracted by the CNN layer in time order and effectively overcome the long-term dependence of traditional RNN on time series data. In addition, the global attention mechanism of Luong attention is introduced to assign different attention weights to all hidden layers of BiGRU, thus strengthening the influence of important feature information on the current output value. The use of dropout layers can mask some neurons, thereby preventing the model from overfitting and enhancing the model's generalization ability. According to the multi-task learning framework, the initial SOC and SOE values are output through two fully connected layers under the mechanism of information sharing and complementarity between multiple tasks. The bidirectional filtering strategy of SREKS is used to smooth the initial estimation. SREKS first uses the SREKF algorithm for forward filtering, then uses a fixed interval grinding technology for secondary filtering on the basis of forward filtering. Finally, the bidirectional filtered values are output to obtain high-precision estimated SOC and SOE values.

3. Experimental analysis

3.1. Lithium-ion battery test platform design

The test platform of this experiment mainly includes a temperature laboratory, charge and discharge test, and BMS. A ternary lithium-ion battery with a standard capacity of 72Ah is selected for lithium-ion batteries. The test platform and related parameters of the ternary lithium-ion battery are shown in Fig. 9.

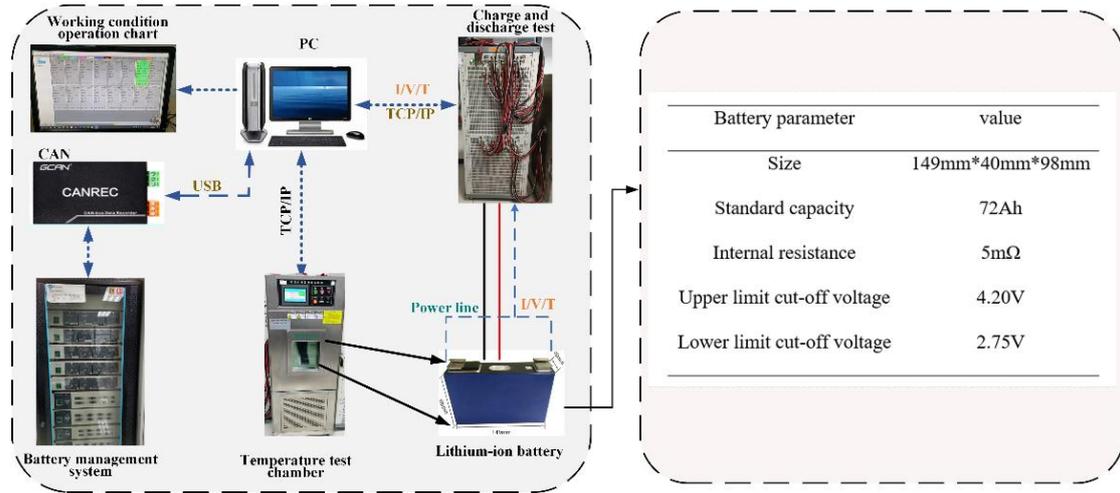


Fig. 9. The experimental platform

This study carried out the HPPC and BBDST tests of ternary lithium-ion battery at different temperature environments. HPPC and BBDST test data at room temperature (25°C) are used as training data, the training data are shown in Fig. 10.

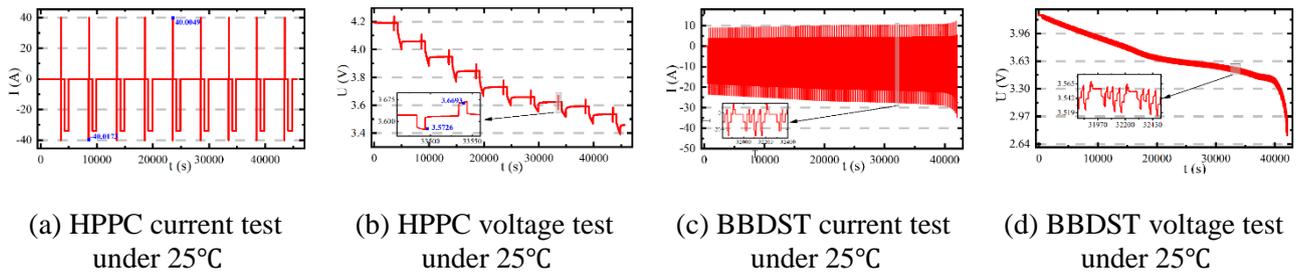
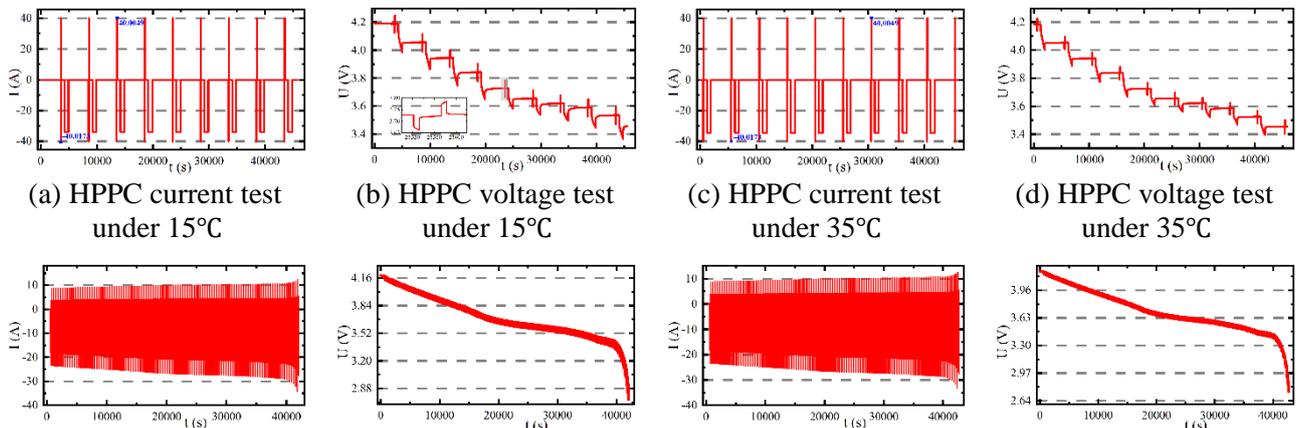


Fig. 10. Training data

HPPC and BBDST test data at 15°C and 35°C are used to verify the performance of the AMC-BiGRU-SREKS model. To better evaluate the range of new-energy vehicles in the case of daily temperature changes to ensure driving safety. The test data are shown in Fig. 11.



(e) BBDST current test under 15°C

(f) BBDST voltage test under 15°C

(g) BBDST current test under 35°C

(h) BBDST voltage test under 35°C

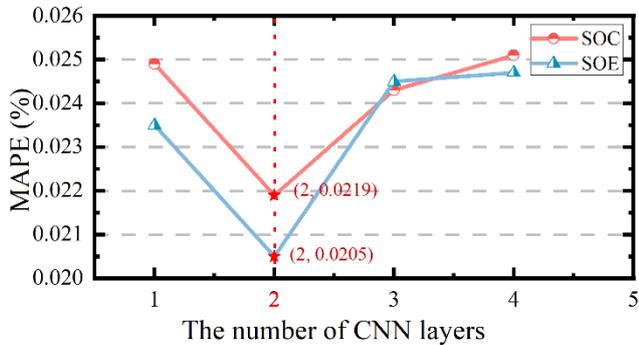
Fig.11. Test data

3.2. Hyperparameter setting

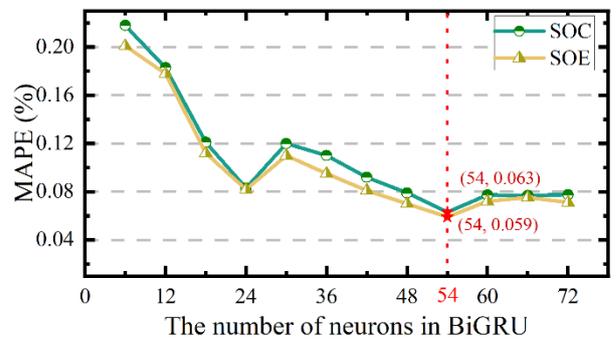
To obtain the optimal model, this paper uses control variables to optimize the hyperparameters. To better reflect the prediction effect of the algorithms, MAPE is used to reflect the accuracy of the model, MAE is used to reflect the actual situation of the prediction error, and RMSE is used to reflect the sample standard deviation of the difference between the predicted and actual value. Their calculation formulas are shown in Equation (10).

$$\begin{cases} MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \\ MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \\ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \end{cases} \quad (10)$$

Where y_i is the real value and \hat{y}_i is the estimated value. To get a more accurate model, use MAPE as a performance indicator during the tuning process. First, set initial values for some parameters. Suppose the number of iterations is 1, the dropout rate is 0.2, and the batch size is 200. Then, using the control variable method to adjust the hyperparameters of the CNN layer, BiGRU neurons, batch size, and learning rate. The results are shown in Fig. 12.



(a) CNN layer number adjustment



(b) BiGRU Neuronal adjustment

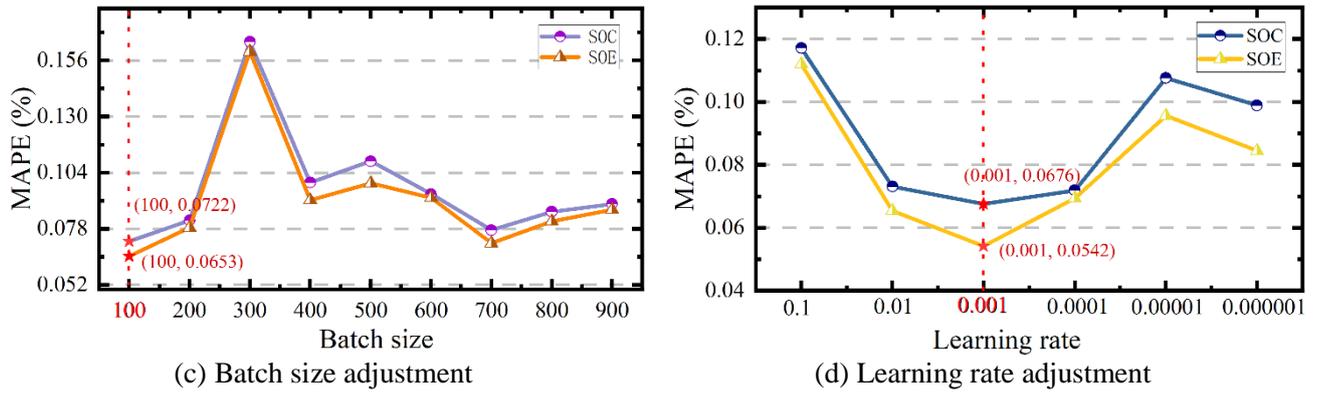


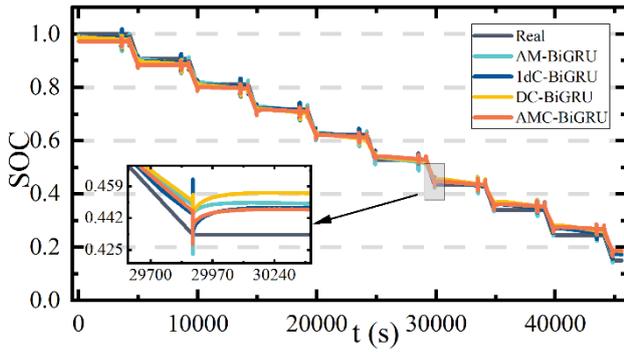
Fig. 12. Process of hyperparameter adjustment

According to Fig. 11, when the number of CNN layers is 2, the number of neurons in BiGRU's hidden layer is 54, the batch size is 100 and the learning rate is 0.001, the MAPE value is lower. The lower the MAPE value, the higher the model's precision. Two convolutional layers in CNN set the convolution kernel size to 3 and filter length to 64. To obtain the critical information on SOC and SOE, two attention layers are used, and the dropout layer is introduced to prevent the occurrence of overfitting.

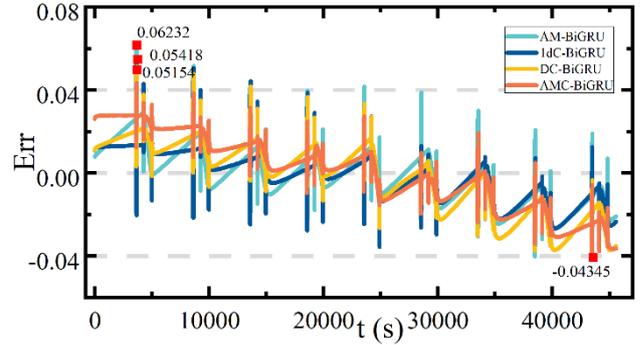
3.3. Prediction results

3.3.1. SOC and SOE prediction with the AMC-BiGRU network

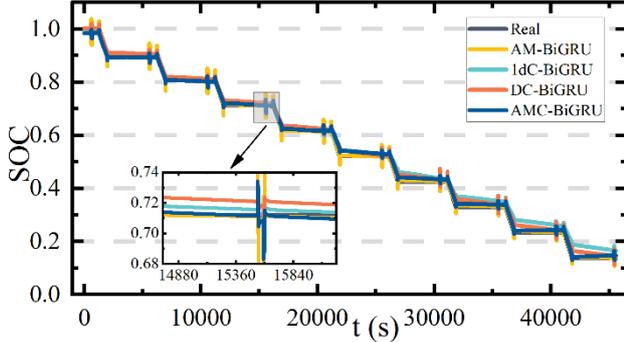
The AMC-BiGRU network uses HPPC and BBDST test data at room temperature (25°C) as training data and uses HPPC and BBDST test data at 15°C and 35°C to evaluate the network performance. In order to verify the accuracy of the multi-task prediction model based on the AMC-BiGRU algorithm constructed in this paper. Based on the research of lithium-ion battery state prediction, this paper constructed the attention mechanism and the bidirectional gate recurrent unit (AM-BiGRU) model based on reference [46], dilated convolutional neural network and bidirectional gate recurrent unit (DC-BiGRU) model based on reference [47] and one-dimensional convolutional neural network and bidirectional gate recurrent unit (1dC-BiGRU) model based on reference [48]. The estimation accuracy and performance of different optimized BiGRU models are compared. The predicted results of SOC using HPPC test data at 25°C as training data are shown in Fig. 13.



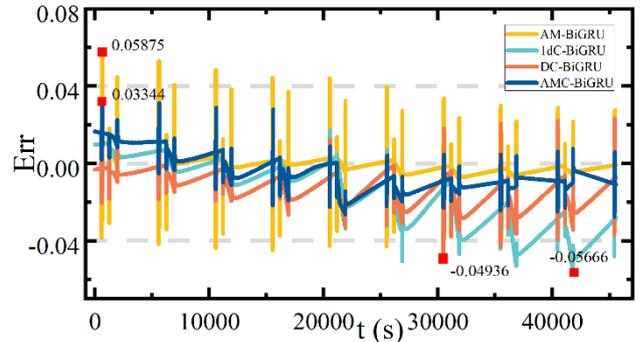
(a) SOC estimation results under HPPC at 15°C



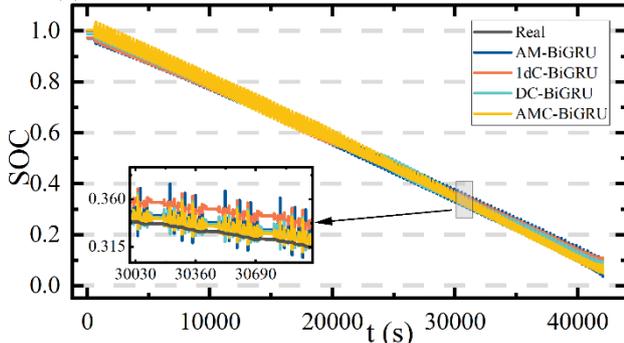
(b) SOC estimation error at 15°C HPPC



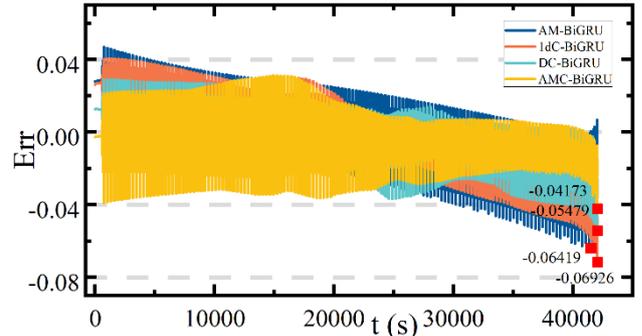
(c) SOC estimation results under HPPC at 35°C



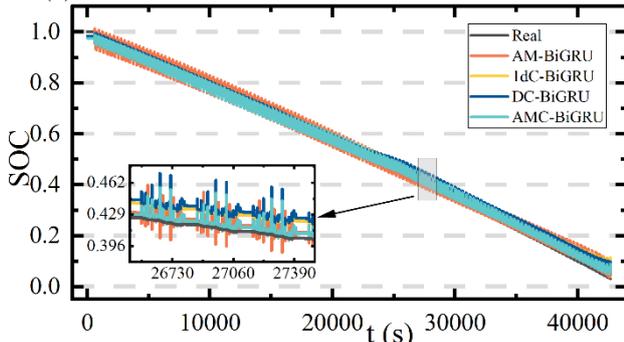
(d) SOC estimation error at 35°C HPPC



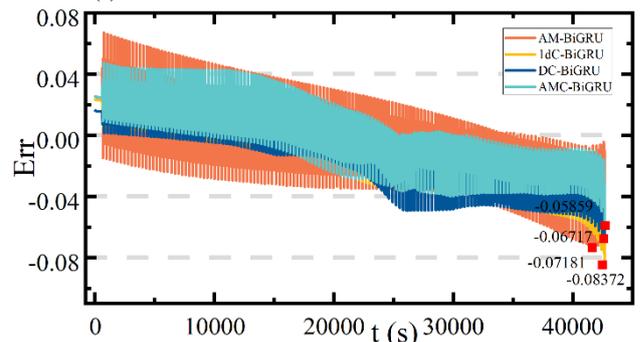
(e) SOC estimation results under BBDST at 15°C



(f) SOC estimation error at 15°C BBDST



(g) SOC estimation results under BBDST at 35°C



(h) SOC estimation error at 35°C BBDST

Fig. 13. SOC prediction results under complex HPPC and BBDST conditions

According to Fig. 13, the maximum absolute error of the AMC-BiGRU to predict SOC under HPPC conditions at 15°C and 35°C is within 0.04345 and 0.03344, respectively. Meanwhile, the maximum absolute error of SOC under BBDST conditions at 15°C and 35°C is within 0.04173 and 0.05859, respectively. Compared with the AM-BiGRU,

1dC-BiGRU, and DC-BiGRU algorithms, the AMC-BiGRU has a minor error value, so the prediction accuracy is higher. The predicted results of SOE using HPPC test data at 25°C as training data are shown in Fig. 14.

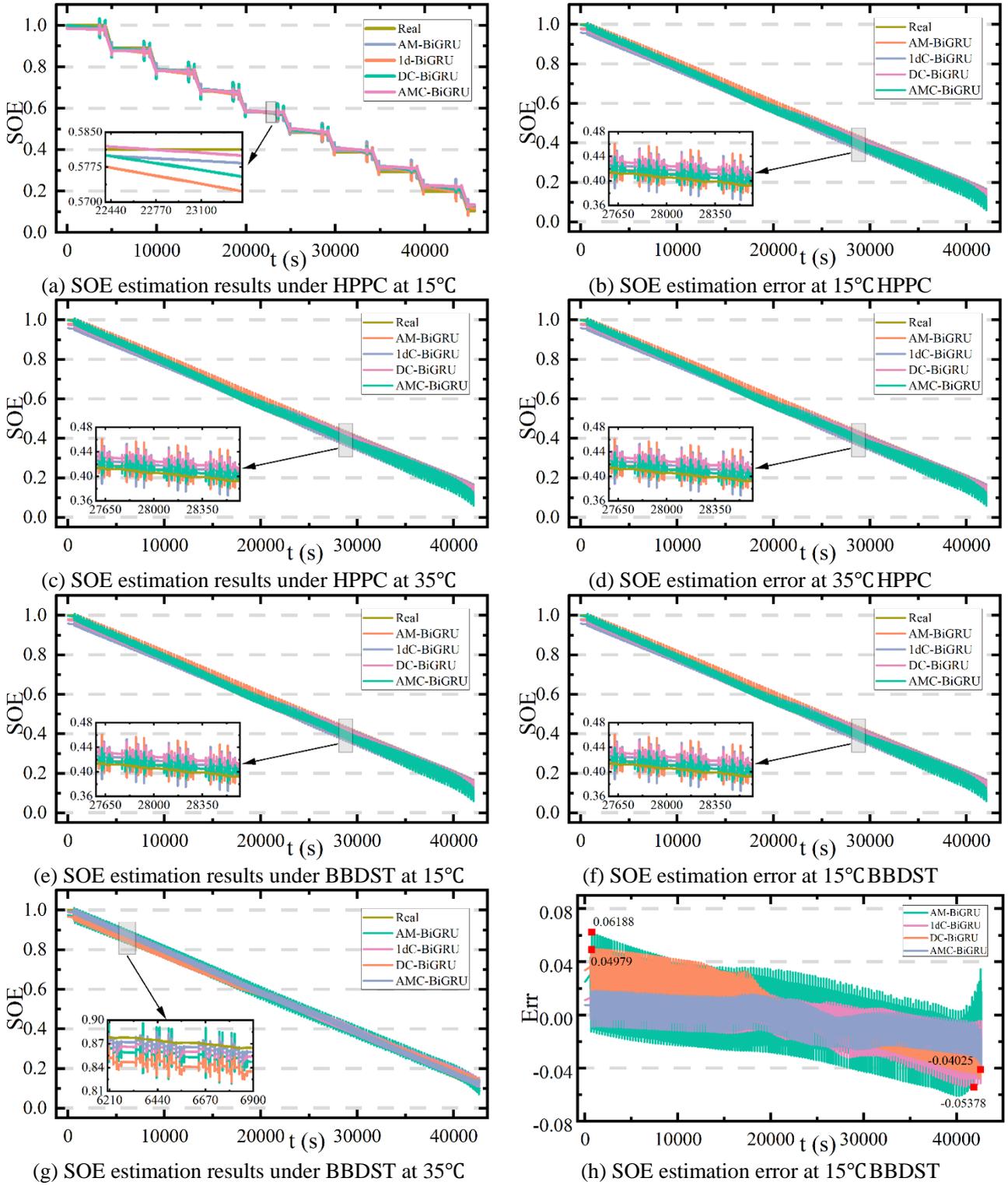


Fig. 14. SOE prediction results under complex HPPC and BBDST conditions

Fig. 14 shows that AM-BiGRU, 1dC-BiGRU, and DC-BiGRU algorithms all have high prediction errors. On the

contrary, the SOE prediction results of the AMC-BiGRU have a high degree of fitting with the desired value. The maximum absolute error of the AMC-BiGRU model to predict SOE at 15°C and 35°C is within 0.03896 and 0.04170, respectively. Meanwhile, the maximum absolute error of SOC under BBDST conditions at 15°C and 35°C is within 0.05651 and 0.04025, respectively. In this paper, MAE, MAPE, and RMSE are used to reflect the predictive performance of the constructed model, and the calculation results are shown in Fig. 15.

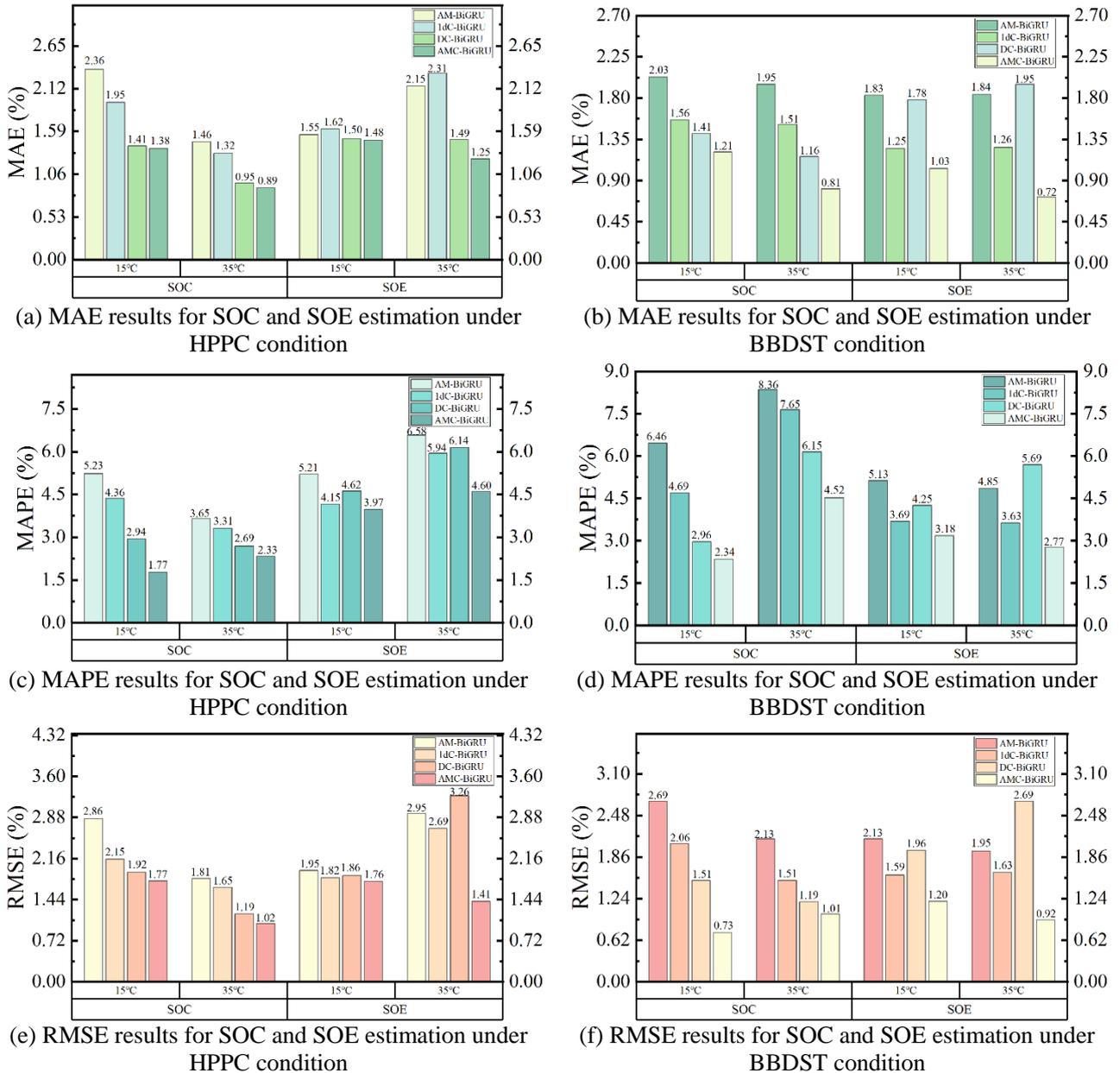


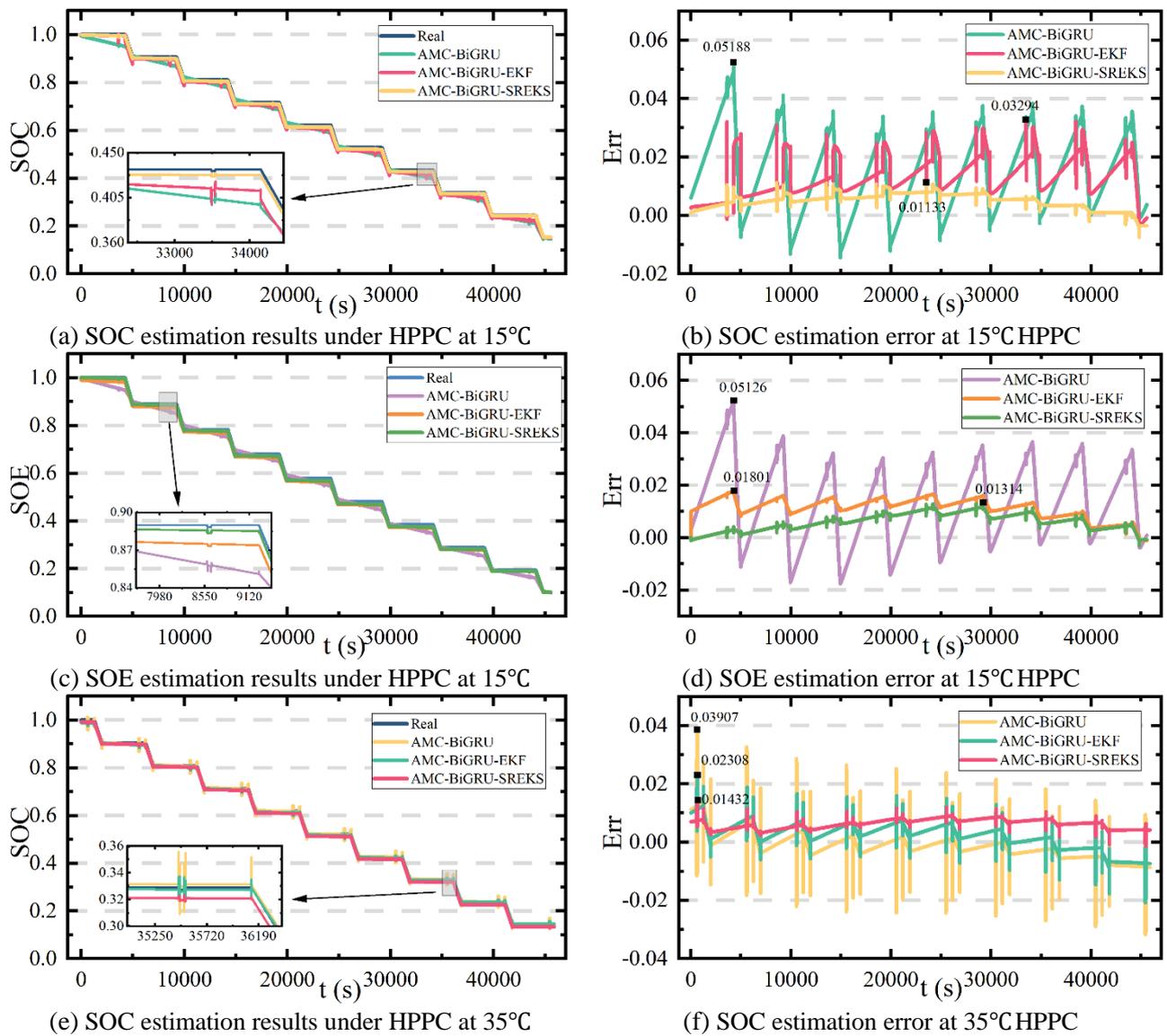
Fig. 15. MAE, MAPE, and RMSE results of the four algorithms under HPPC and BBDST conditions

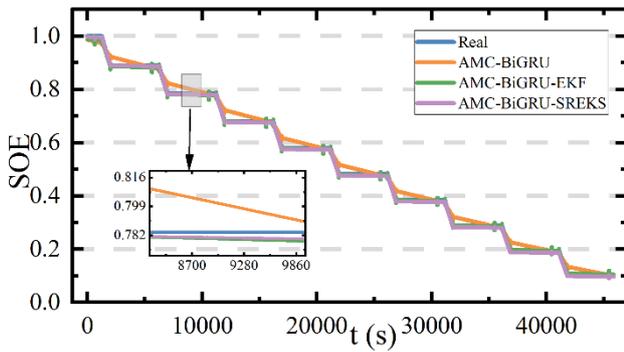
In Fig. 15, the AMC-BiGRU algorithm has low MAE, MAPE, and RMSE values, which are within 1.48%,

4.60%, and 1.77%. This can be seen from the experimental results that AMC-BiGRU has higher prediction accuracy than AM-BiGRU, 1dC-BiGRU, and DC-BiGRU.

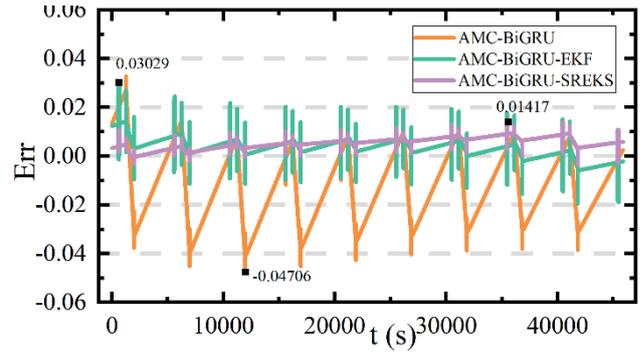
3.3.2. SOC and SOE prediction after SREKS filtering

Based on AMC-BiGRU, the SREKS algorithm is proposed to form the AMC-BiGRU-SREKS algorithm. The output of AMC-BiGRU is taken as the input of SREKS and smoothed it. BBDST test data at room temperature (25°C) are used as training data, and HPPC test data at 15°C and 35°C are used to verify the performance of the AMC-BiGRU-SREKS algorithm. The predicted results are shown in Fig. 16.





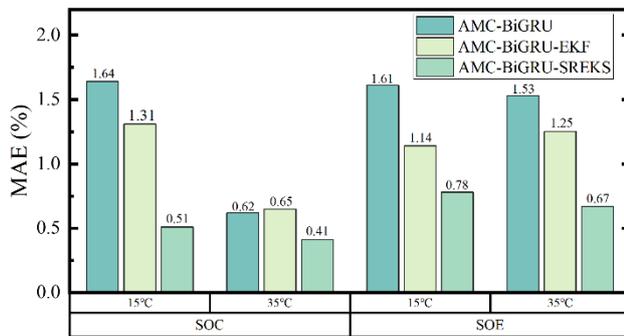
(g) SOE estimation results under HPPC at 35°C



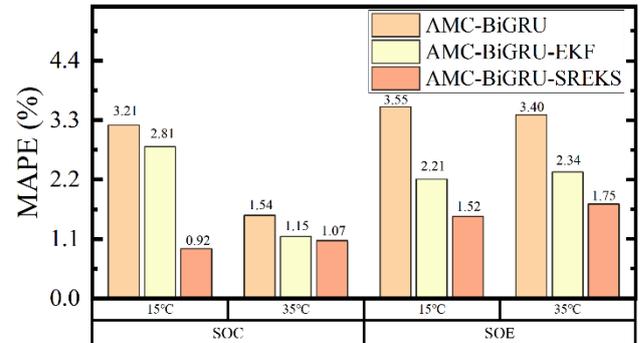
(h) SOE estimation error at 35°C HPPC

Fig. 16. SOC and SOE prediction results for HPPC at 15°C and 35°C

Based on the above model, the AMC-BiGRU-EKF algorithm is constructed to confirm the high accuracy of the AMC-BiGRU-SREKS algorithm. As shown in Fig. 16, compared with AMC-BiGRU and AMC-BiGRU-EKF algorithms, the SOC and SOE estimated by the AMC-BiGRU-SREKS method have the lowest error values, 0.01133, and 0.01134, respectively, under the HPPC conditions at 15°C. At the same time, the error values of SOC and SOE estimated by the AMC-BiGRU-SREKS at 35°C are also the lowest, which are 0.01432 and 0.01417, respectively. To further reflect the performance of several algorithm, the MAE and MAPE performance parameters of the three algorithms are obtained and compared, and the results are shown in Fig. 17.



(a) MAE results for SOC and SOE estimation



(b) MAPE results for SOC and SOE estimation

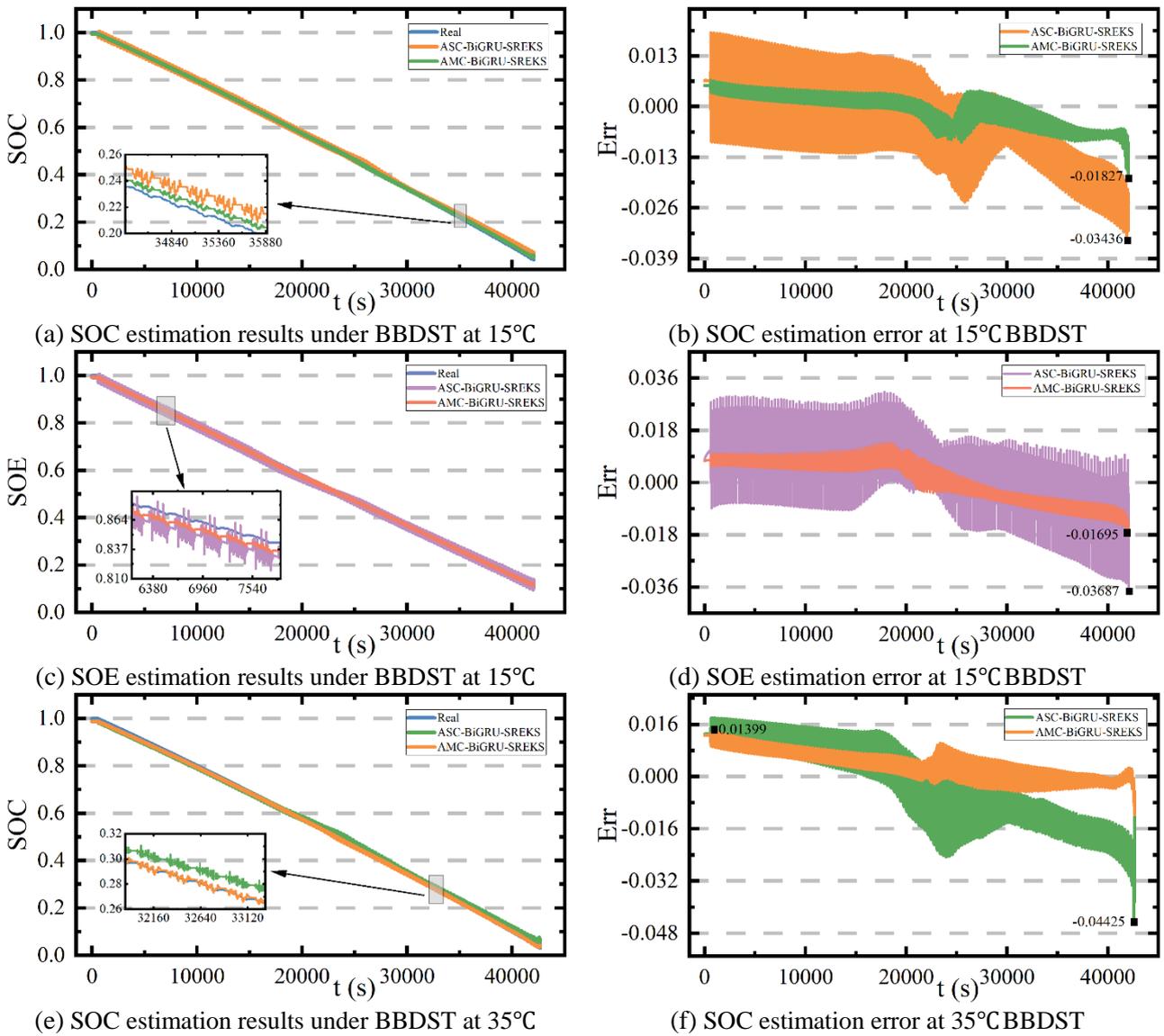
Fig. 17. MAE and MAPE results of the algorithms under HPPC working conditions

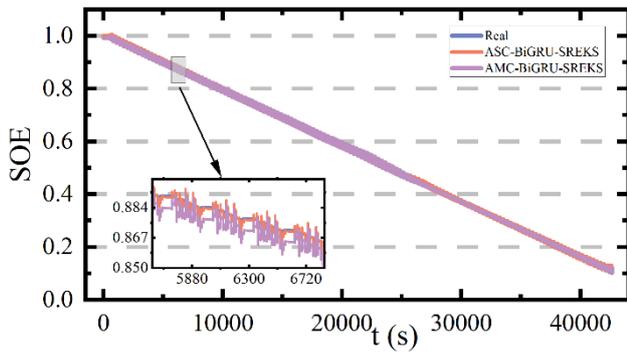
In Fig. 17, the results of SOC and SOE using the AMC-BiGRU-SREKS method have lower MAE and MAPE values at two different temperatures. Thus, it can be proved that the algorithm has a high degree of fitting, and the estimates are close to ideal values.

3.3.3. Comparison with a single-task learning model

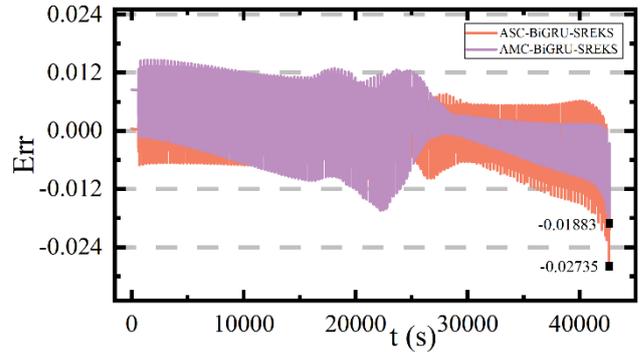
Compared with the STL algorithm, MTL obtains more information and improves output accuracy because of task coupling. Due to MTL output-related tasks at the same time, it greatly shortens the time of prediction as well as improves the efficiency of prediction. The experimental comparison is conducted between single-task (ASC-BiGRU-SREKS) and multi-task (AMC-BiGRU-SREKS). HPPC test data at room temperature (25°C) are used as training data, and BBDST test data at 15°C and 35°C are used to verify the algorithm's performance, as shown in Fig. 18.

Fig. 18.





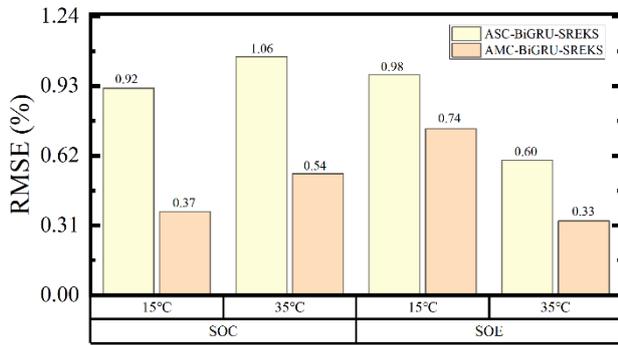
(g) SOE estimation results under BBDST at 35°C



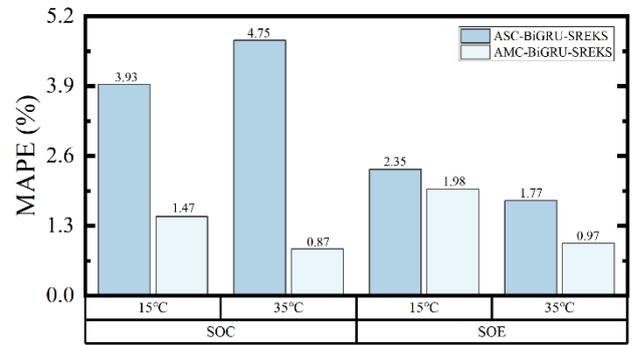
(h) SOE estimation error at 35°C BBDST

Fig. 18. SOC and SOE prediction results for BBDST at 15 °C condition and 35 °C

In Fig. 18, compared with the STL, the SOC, and SOE estimated by the MTL algorithm at 15 °C have the lowest absolute error values, which are 0.01827 and 0.01695, respectively. Meanwhile, the absolute error of SOC and SOE estimated by the MTL at 35 °C BBDST are 0.01399 and 0.01883, respectively. The RMSE and MAPE performance parameters of the two algorithms are obtained and compared, and the calculation results are shown in Fig. 19.



(a) RMSE



(b) MAPE

Fig. 19. RMSE and MAPE results of the algorithms under BBDST working conditions

It is evident from Fig. 19 that compared with the single-tasking algorithm, the multi-task learning mechanism has smaller RMSE and MAPE values, thus a better fitting degree.

4. Conclusion

In this paper, a multi-task learning algorithm based on the attention mechanism of the convolutional neural network-bidirectional gate recurrent unit and the square root extended Kalman smoothing (AMC-BiGRU-SREKS) is proposed to predict the SOC and SOE of ternary lithium-ion batteries under different complex

conditions. The deep learning is combined with the filter algorithm, which solves the long time, low precision, and strong nonlinear in the SOC and SOE estimation of lithium-ion batteries by the traditional algorithm. Thus, it improves the safety of the BMS and ensures the safety of new-energy vehicles. Firstly, the CNN-BiGRU model is constructed. CNN is used to capture the spatial characteristics of the input data, and BiGRU is used to capture the long-term dependencies of the time series data. Secondly, AM and MTL mechanisms are added to strengthen the influence of essential feature parameters on output. The hard sharing of parameters is used to obtain the output of multiple tasks simultaneously. Finally, the SREKS algorithm is introduced to carry out smooth bidirectional filtering on the output results of the neural network model and constantly improve the output accuracy of SOC and SOE. In addition, the algorithm's performance is evaluated under HPPC and BBDST conditions at 15□, 25□, and 35□. Under HPPC conditions, the MAE and MAPE of SOC are within 0.51% and 1.07%, and the MAE and MAPE of SOE are within 0.78% and 1.75%, respectively. Under BBDST conditions, the RMSE and MAPE of SOC are within 0.54% and 1.47%, and the RMSE and MAPE of SOE are within 0.74% and 1.98%, respectively. Experimental results show that the algorithm has high precision and robustness.

Acknowledgments

The work was supported by the National Natural Science Foundation of China (No. 62173281, 61801407), Sichuan Science and Technology Program (No. 23ZDYF0734, 23NSFSC4444), Dazhou City School Cooperation Project (No. DZXQHZ006), Technopole Talent Summit Project (No. KJCRCFH08), and Robert Gordon University.

References

1. Zhang, Z., et al., *A Nonlinear Prediction Method of Lithium-Ion Battery Remaining Useful Life Considering Recovery Phenomenon*. International Journal of Electrochemical Science, 2020. **15**(9): p. 8674-8693.
2. Xia, F.J., et al., *Revealing structural degradation in layered structure oxides cathode of lithium ion batteries via in-situ transmission electron microscopy*. Journal of Materials Science & Technology, 2023.

154: p. 189-201.

3. Liu, Y., et al., *A review of lithium-ion battery state of charge estimation based on deep learning: Directions for improvement and future trends*. Journal of Energy Storage, 2022. **52:** p. 1-15.
4. Wu, L., et al., *Online SOC Estimation Based on Simplified Electrochemical Model for Lithium-Ion Batteries Considering Current Bias*. Energies, 2021. **14(17):** p. 5265-5277.
5. Shrivastava, P., et al., *Combined State of Charge and State of Energy Estimation of Lithium-Ion Battery Using Dual Forgetting Factor-Based Adaptive Extended Kalman Filter for Electric Vehicle Applications*. IEEE Transactions on Vehicular Technology, 2021. **70(2):** p. 1200-1215.
6. Shrivastava, P., et al., *Model-based state of X estimation of lithium-ion battery for electric vehicle applications*. International Journal of Energy Research, 2022. **46(8):** p. 10704-10723.
7. Fasahat, M. and M. Manthouri, *State of charge estimation of lithium-ion batteries using hybrid autoencoder and Long Short Term Memory neural networks*. Journal of Power Sources, 2020. **469:** p. 1-8.
8. Ni, Z.C., B. Li, and Y. Yang, *Deep domain adaptation network for transfer learning of state of charge estimation among batteries*. Journal of Energy Storage, 2023. **61:** p. 1-14.
9. Lin, Q., et al., *Stable and Accurate Estimation of SOC Using eXogenous Kalman Filter for Lithium-Ion Batteries*. Sensors, 2023. **23(1):** p. 467-480.
10. Shi, H., et al., *Improved electric-thermal-aging multi-physics domain coupling modeling and identification decoupling of complex kinetic processes based on timescale quantification in lithium-ion batteries*. Applied Energy, 2024. **353:** p. 1-22.
11. Wang, L., et al., *A novel OCV curve reconstruction and update method of lithium-ion batteries at different temperatures based on cloud data*. Energy, 2023. **268:** p. 1-15.

12. Liu, D., et al., *A novel fuzzy-extended Kalman filter-ampere-hour (F-EKF-Ah) algorithm based on improved second-order PNGV model to estimate state of charge of lithium-ion batteries*. International Journal of Circuit Theory and Applications, 2022. **50**(11): p. 3811-3826.
13. Lai, X., et al., *A joint state-of-health and state-of-energy estimation method for lithium-ion batteries through combining the forgetting factor recursive least squares and unscented Kalman filter*. Measurement, 2022. **205**: p. 1-11.
14. Alipouri, Y., S. Zhao, and B. Huang, *Distributed data-driven observer for linear time invariant systems*. International Journal of Adaptive Control and Signal Processing, 2020. **34**(4): p. 503-519.
15. Zhong, J., et al., *Data-driven Crowd Modeling Techniques: A Survey*. Acm Transactions on Modeling and Computer Simulation, 2022. **32**(1): p. 1-15.
16. Ali, O., et al., *On-line WSN SoC estimation using Gaussian Process Regression: An Adaptive Machine Learning Approach*. Alexandria Engineering Journal, 2022. **61**(12): p. 9831-9848.
17. Wang, Y.H., H.H. Huang, and H.X. Wang, *A new method for fast state of charge estimation using retired battery parameters*. Journal of Energy Storage, 2022. **55**: p. 1-9.
18. Wang, S., Y. Ning, and H. Shi, *A new uncertain linear regression model based on equation deformation*. Soft Computing, 2021. **25**(20): p. 12817-12824.
19. Qi, X.Y., et al., *Recent Advance of Machine Learning in Selecting New Materials*. Acta Chimica Sinica, 2023. **81**(2): p. 158-174.
20. Ji, S.L., et al., *Deep learning enhanced lithium-ion battery nonlinear fading prognosis*. Journal of Energy Chemistry, 2023. **78**: p. 565-573.
21. Pant, S., et al., *Enforcing Intelligent Learning-Based Security in Internet of Everything*. Ieee Internet of Things Journal, 2023. **10**(4): p. 3071-3078.

22. Goodfellow, I., et al., *Generative Adversarial Networks*. Communications of the Acm, 2020. **63**(11): p. 139-144.
23. Hourcade, C., M. Bonnin, and E. Beucler, *New CNN-based tool to discriminate anthropogenic from natural low magnitude seismic events*. Geophysical Journal International, 2023. **232**(3): p. 2119-2132.
24. Li, Y.F., et al., *DCT-GAN: Dilated Convolutional Transformer-Based GAN for Time Series Anomaly Detection*. Ieee Transactions on Knowledge and Data Engineering, 2023. **35**(4): p. 3632-3644.
25. Zou, R.M., et al., *A novel convolutional informer network for deterministic and probabilistic state-of-charge estimation of lithium-ion batteries*. Journal of Energy Storage, 2023. **57**: p. 1-19.
26. Zhang, Z., et al., *An Improved Bidirectional Gated Recurrent Unit Method for Accurate State-of-Charge Estimation*. Ieee Access, 2021. **9**: p. 11252-11263.
27. Bian, C., H. He, and S. Yang, *Stacked bidirectional long short-term memory networks for state-of-charge estimation of lithium-ion batteries*. Energy, 2020. **191**: p. 1-10.
28. Wang, S., et al., *Improved singular filtering-Gaussian process regression-long short-term memory model for whole-life-cycle remaining capacity estimation of lithium-ion batteries adaptive to fast aging and multi-current variations*. Energy, 2023. **284**: p. 1-14.
29. Fan, T.-E., et al., *Simultaneously estimating two battery states by combining a long short-term memory network with an adaptive unscented Kalman filter*. Journal of Energy Storage, 2022. **50**: p. 1-12.
30. Ma, L., C. Hu, and F. Cheng, *State of Charge and State of Energy Estimation for Lithium-Ion Batteries Based on a Long Short-Term Memory Neural Network*. Journal of Energy Storage, 2021. **37**: p. 1-10.
31. Hong, J.-K., *Vibration Prediction of Flying IoT Based on LSTM and GRU*. Electronics, 2022. **11**(7): p. 1-15.
32. Busari, G.A. and D.H. Lim, *Crude oil price prediction: A comparison between AdaBoost-LSTM and*

- AdaBoost-GRU for improving forecasting performance*. Computers & Chemical Engineering, 2021. **155**: p. 1-9.
33. Hansun, S., A. Wicaksana, and A.Q.M. Khaliq, *Multivariate cryptocurrency prediction: comparative analysis of three recurrent neural networks approaches*. Journal of Big Data, 2022. **9**: p. 1-16.
34. Zhao, J., et al., *Prediction model for stock price trend based on recurrent neural network*. Journal of Ambient Intelligence and Humanized Computing, 2021. **12**(1): p. 745-753.
35. Wang, D.H., et al., *Neural Network-Based State of Charge Estimation Method for Lithium-ion Batteries Based on Temperature*. Intelligent Automation and Soft Computing, 2023. **36**(2): p. 2025-2040.
36. Dong, Z., et al., *Internal Cascaded Neuromorphic Computing System for Fast Electric Vehicle State of Charge Estimation*. IEEE Transactions on Consumer Electronics, 2023: p. 1-5.
37. Yang, X., et al., *A novel fuzzy adaptive cubature Kalman filtering method for the state of charge and state of energy co-estimation of lithium-ion batteries*. Electrochimica Acta, 2022. **415**: p. 1-13.
38. Zhang, S.Z. and X.W. Zhang, *A novel non-experiment-based reconstruction method for the relationship between open-circuit-voltage and state-of-charge/state-of-energy of lithium-ion battery*. Electrochimica Acta, 2022. **403**: p. 1-19.
39. Shi, H., et al., *A Novel Dual Correction Extended Kalman Filtering Algorithm for The State of Charge Real-Time Estimation of Packing Lithium-Ion Batteries*. International Journal of Electrochemical Science, 2020. **15**(12): p. 12706-12723.
40. Tian, J., et al., *Capacity attenuation mechanism modeling and health assessment of lithium-ion batteries*. Energy, 2021. **221**: p. 1-14.
41. Fan, T.-E., et al., *Simultaneously estimating two battery states by combining a long short-term memory network with an adaptive unscented Kalman filter*. Journal of Energy Storage, 2022. **50**: p. 1-14.

42. Ma, L., C. Hu, and F. Cheng, *State of Charge and State of Energy Estimation for Lithium-Ion Batteries Based on a Long Short-Term Memory Neural Network*. Journal of Energy Storage, 2021. **37**: p. 1-19.
43. Wu, F., et al., *An improved long short-term memory based on global optimization square root extended Kalman smoothing algorithm for collaborative state of charge and state of energy estimation of lithium-ion batteries*. International Journal of Circuit Theory and Applications, 2023. **51**(8): p. 3880-3896.
44. Pang, H., et al., *A Composite State of Charge Estimation for Electric Vehicle Lithium-Ion Batteries Using Back-Propagation Neural Network and Extended Kalman Particle Filter*. Journal of the Electrochemical Society, 2022. **169**(11): p. 516-524.
45. Wang, Y.J., et al., *A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems*. Renewable & Sustainable Energy Reviews, 2020. **131**: p. 1-18.
46. Liu, P., et al., *A Data-Driven Comprehensive Battery SOH Evaluation and Prediction Method Based on Improved CRITIC-GRA and Att-BiGRU*. Sustainability, 2023. **15**(20): p. 1-15.
47. Bao, Z., et al., *A New Hybrid Neural Network Method for State-of-Health Estimation of Lithium-Ion Battery*. Energies, 2022. **15**(12): p. 1-16.
48. Mazzi, Y., H. Ben Sassi, and F. Errahimi, *Lithium-ion battery state of health estimation using a hybrid model based on a convolutional neural network and bidirectional gated recurrent unit*. Engineering Applications of Artificial Intelligence, 2024. **127**: p. 1-13.